2019 IMAG MSM Meeting - Abstract Submission Form
2019 Multiscale Modeling Consortium Meeting - Translation and Dissemination (March 6-7, 2019)

PI(s) of MSM U01: N/A

Institution(s): Jacob Barhak for Poster 1 / Jacob Barhak and Joshua Schertz for Poster 2

MSM U01 Grant Number: N/A

Please note that those are twin related posters that will fit the space allocated for one poster. The posters are interactive and can be accessed from the link:
https://jacob-barhak.github.io/InteractivePoster_MSM_IMAG_2019.html

The poster abstracts are in the next few pages.

POSTER 1

Title: The Reference Model is the most validated diabetes cardiovascular model known

Abstract Authors

Jacob Barhak

Abstract Text

The Reference Model is an ensemble model that accumulates knowledge from multiple other models and validates this knowledge against multiple populations. After connecting to ClinicalTrials.Gov it has been growing rapidly and has reached the point where it validates against more populations than any other known diabetes model. It currently contains 30 risk models that cooperate and compete and assemble the best model that fits 123 cohorts from 31 populations. This year there was an increase in the number of cohort downloaded from ClinicalTrials.Gov, yet more importantly the cumulative computational gap of knowledge can now be explored interactively via the web. This gap of knowledge shows the difference between the model predication and the results for each clinical trial cohort. The Reference Model accumulates models and data, so being able to show this gap for accumulated knowledge represents our limits to model diabetes. If ClinicalTrials.Gov would be better standardized, it would be even easier to import data. More data can help narrow this gap that can now be calculated and visualized. Such improvements may lead in the long run for better models that may be used for decision making now reserved for humans. However, to allow such advanced modeling, ClinicalTrials.Gov data requires standardization.

Context: This poster described advances since the poster last year. It is one poster from two poster. The other poster describes efforts to improve ClinicalTrials.Gov standardization to help importing data.

Data: The model is based on published risk equations and on data extracted from the literature and from ClinicalTrials.Gov.

Evaluation: The Reference Model is a validation model - it assembled and validates multiple models against actual results recorded in clinical trials.

Limitations: The model is currently limited by the availability to extract data and the rate it is added to ClinicalTrials.Gov . It is also limited by availability of computing power since simulations take over 2 years on a single CPU.

Version control: The Reference Model code versions and results are archived and backed up privately since it is proprietary, Some versions of the MIcro Simulation Tool, that is the engine that executes the model in HPC environment, are partially released on GitHub while more advanced versions are kept private and not released.

Documentation: The Reference Model technologies are documented in many publications, many in MODSIM and MSM/IMAG meetings. Its development story is kept for over 6 years of development.

Dissemination: Currently only techniques are disseminated alongside results. Most techniques are available in USPTO as a US patent and pending patent.

Independent reviews: Multiple publications of the techniques were scrutinized. The Reference Model performance was compared to other models in modeling challenges. The Reference Model now replaces the role of challenges since it validates multiple models. It now provides independent review for multiple other models by validating them all against multiple populations.

Test competing implementations: The Reference Model makes multiple models compete and tests their performance.

Conform to standards: The Reference Model uses XML schema defined by ClinicalTrials.Gov to import data. However, the standardization level of ClinicalTrials.Gov need improvement as discussed in the twin poster.

POSTER 2

Title: Clinical Unit Mapping for Standardization of ClinicalTrials.Gov

Abstract Authors

Jacob Barhak, Joshua Schertz

Abstract Text

ClinicalTrials.Gov now accumulates information from over quarter of a million trials with over 10% recording trial results. It is now a U.S. Law to upload many clinical trials to this fast growing database. Data from this database can be extracted in XML format and used for modeling. The Reference Model, for example, extracts population baseline statistics and trial results. However, the database is based on textual input and although scrutinized by humans, it is currently designed for human comprehension rather than computer comprehension. Specifically, non standardized units prevent machine comprehending associated numbers. On 7 Feb 2019 all 34,751 trials with results were downloaded and unit fields were indexed and analyzed. There were 23,733 different units detected. This is a clear sign that standardization is required. We used some machine learning and Natural Language Processing algorithms to organize the data for easier processing by humans. We then created the web site: ClinicalUnitMapping.com to help standardize the units so that many models can process data in this valuable database. If units are standardized, the valuable numerical data in this database can become machine comprehensible.

Context: This poster described data preparation towards standardization work of units within ClinicalTrials.Gov

Data: The units and trial data were extracted from ClinicalTrials.Gov, auxiliary units for standardization were extracted from CDISC.

Evaluation: N/A

Limitations: N/A

Version control: Code versions are kept locally, the web site code is held in a private Github repository.

Documentation: The ClinicalUnitMapping.com web site contains an about page with publications.

Dissemination: This standardization work was presented in SISO and in the COMBINE meeting. The web site is public.

Independent reviews: N/A

Test competing implementations: N/A

Conform to standards: The unit mapping for this web site refers to existing CDISC/UCUM specifications. The intention is that future work will contribute to CDISC and eventually be imported to UMLS.